Theoretical study of structural changes in DNA under high external hydrostatic pressure

P Sudheer Kumar

A thesis submitted to

Indian Institute of Science Education and Research, Pune

In partial fulfilment of the requirements for the

BS-MS Dual Degree Programme

Thesis Supervisor: Dr Anirban Hazra

April, 2014

Indian Institute of Science Education and Research, Pune

Dr Homi Bhabha Road, Pashan, Pune India 411008

Certificate

This is to certify that this dissertation entitled "Theoretical study of structural changes in DNA under high external hydrostatic pressure" towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents original research carried out by "P Sudheer Kumar" at "Indian Institute of Science Education and Research, Pune" under the supervision of "Dr Anirban Hazra, Assistant Professor, Department of Chemistry" during the academic year 2013-2014.

Coordinator, Coordinator of Chemistry Faculty

Committee:

Dr Anirban Hazra

Reader 1

Reader 2

Declaration

I hereby declare that the matter embodied in the report entitled "Theoretical study of structural changes in DNA under high external hydrostatic pressure" are the results of the investigations carried out by me at the Department of Chemistry, Indian Institute of Science Education and Research, Pune under the supervision of Dr Anirban Hazra and the same has not been submitted elsewhere for any other degree.

Project Supervisor

P Sudheer Kumar

Dr Anirban Hazra

Acknowledgements

I would like to thank my project supervisor Dr Anirban Hazra for giving me freedom in pursuing my own problem and all his support throughout my project with him.

Also I would like to thank my lab mates Mahesh, Avdhoot and Meghna. I enjoyed working with all of them and was happy in participating in their discussions which were very stimulating. I would like to thank Mahesh for teaching me GAMESS program package and quantum calculations.

I am very grateful towards Dr Arnab Mukherjee and his students for teaching and helping me in learning molecular dynamics simulations and GROMACS package.

Contents

List of figures

Abstract

Structural changes of the DNA macromolecule under high external hydrostatic pressure (2000 bar) have been studied using quantum chemical calculations and molecular dynamics simulations. Such studies give insight into the governing forces in DNA, helps understand the importance of hydration layers, and acts as a basis for understanding the biology of organisms living under high pressure. The structural changes calculated from quantum calculations correspond to small compressions in the hydrogen bond lengths (around 0.016 Å in A-T base pair and 0.012 Å in G-C base pair) and suggest that the DNA molecule is only slightly perturbed due to high external pressure.

Chapter 1

Introduction

Deoxyribonucleic acid (DNA) is the biological storage bank for genetic information. This genetic information is essential in the development and functioning of living organisms. DNA molecules are made up of phosphate groups, sugars (deoxyribose) and nucleobases as shown in Figure 1. The phosphate groups and sugars make up the backbone of the DNA to which the nucleobases are attached. The nucleobases are of two types: the purines (Adenine (A) and Guanine (G)), and the pyrimidines (Thymine (T) and Cytosine (C)). In RNA (ribonucleic acid) another pyrimidine nucleobase called Uracil (U) replaces Thymine. The purines pair up with their complementary pyrimidines in a process called base pairing. Adenine forms two hydrogen bonds with Thymine (complementary to adenine), whereas Guanine forms three hydrogen bonds with Cytosine (complementary to guanine). The moiety containing nucleobases attached with the backbone (phosphate group and sugar) are called nucleotides.

Nucleotides forms the twin strands of the DNA (complementary to each other). In the helical structure of the DNA, the backbones are closer together on one side of the helix than the other side. The part where the backbones are closer is called minor groove and the part where the backbones are far apart is called major groove. These grooves twist around the helical strands of the DNA molecule on opposite sides and can be considered as spirals which run in parallel with phosphate back bone of the DNA.

The DNA helix is stabilized by the hydrogen bonds between the nucleobases and also by the stacking interactions between the aromatic rings of the nucleobases. In the most common DNA (B-DNA), the planes of the nucleobases are aligned

perpendicularly to the axis of the DNA molecule. DNA exists in different possible conformations which are termed as A-DNA, B-DNA, and Z-DNA. Both B-DNA and A-DNA have right handed spiral but compared to B-DNA, the A-DNA has a wider right handed spiral, with a shallow and wide minor groove and a narrow and deep major groove. Whereas, the Z-DNA has left handed spiral (opposite to that of B-DNA and A-DNA).

Only B-DNA and Z-DNA have been seen in living organisms. Hydration, DNA sequence, chemical modifications of nucleobases, surrounding environment of the DNA are the key factors responsible for these different conformations seen in DNA.

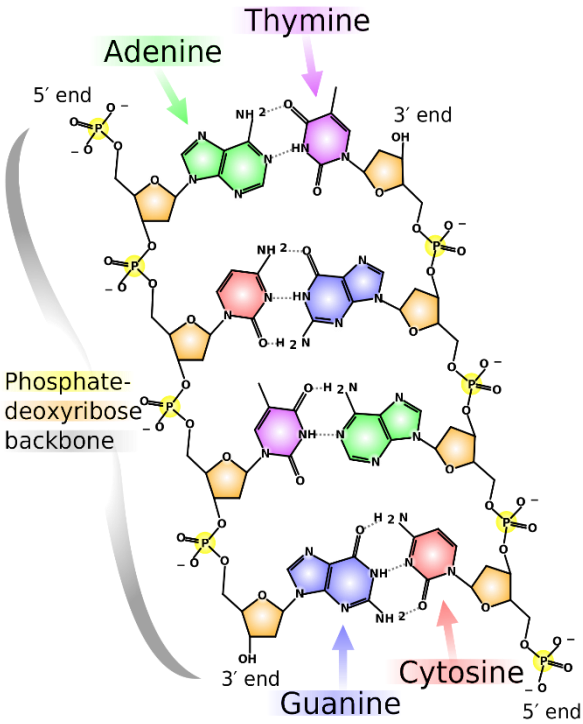


Figure 1: Cartoon representation of the DNA and its corresponding base pairs; A-T and G-C. (*http://en.wikipedia.org/wiki/File:DNA_chemical_structure.svg*)

DNA can witness structural distortions or mutations by various methods which could be chemical or physical in nature. Chemical methods include enzyme activity leading to uncoiling the DNA or breaking the hydrogen bonds, and mutagens changing the chemical composition of DNA. Physical methods can be heating or annealing of the DNA which can easily break the hydrogen bonds between the base pairs and

separate the twin strands of the DNA, high external pressure which can exert a large enough force on the DNA molecule to distort the DNA helical structure.

In deep sea organisms, the DNA along with the other parts of the cell are under high external hydrostatic pressure. Moreover, many such organisms live below the photic zone where there is not enough sunlight for photosynthesis. The other harsh conditions faced by these creatures are very less or no sunlight, food scarcity, and small amounts of oxygen. Because of the depth the pressure in deep oceans ranges from 20 bars to 1100 bars. The pressure increases by about one bar every ten meters.

In 2008, an experimental study on the effects of high external pressure on the structure of a hairpin DNA (which has characteristics of a B-DNA) was reported by Wilton, et al. [1]. It was observed that on application of high external hydrostatic pressure of about 2000 bar (pressures in the order of 1100 bars occur in earth's deep oceans.), the hairpin DNA undergoes slight changes in its structure, which were identified as compressions in the base pairs (0.29 Å in A-T base pair and 0.11 Å in G-C base pair) and slight expansion of the helix in the longitudinal direction due to increase in the size of the grooves. These changes were computed by using ^{1}H NMR chemical shift values of DNA under high external hydrostatic pressure (2000 bar). This calculation was done by using a protocol developed previously which correlates hydrogen bond lengths with the ^{1}H NMR chemical shifts in proteins [2, 3].

The goal of my research is to study structural changes in DNA (more specifically B-DNA which is the most common DNA) under high external hydrostatic pressure, by using accurate quantum chemical calculations and molecular dynamics simulations. The other part is to compare the obtained results between the two theoretical approaches and experiments and understand the specific changes occurring in the DNA structure.

The rest of this thesis will explain the two theoretical methods used in this study. *Chapter 2* gives a brief explanation about the quantum chemical methods and molecular dynamics simulations used in this study, the assumptions and theory considered in the creation of the DNA model to imitate the real DNA. *Chapter 3* provides the results obtained from quantum chemical methods and also from molecular dynamics simulations and a brief discussion about the results. *Chapter 4*

concludes this thesis with a short comparison between the results obtained during this project and the experimental results and a brief explanation about the outcome of this comparison.

Chapter 2

Theoretical methods

Two kinds of theoretical methods have been used in this project which are quantum chemical calculations and molecular dynamics simulations. Quantum chemical calculations is described in detail in section 2.1, whereas molecular dynamics simulations is explained in detail in section 2.2.

2.1 Quantum chemical calculations

In a quantum chemical calculations, the energy of the molecule or the system is obtained as a function of its geometry. This is useful in predicting the restoring force which the molecule produces when an external pressure is applied on the system to distort it from its equilibrium geometry. The new equilibrium geometry (under high external hydrostatic pressure) is obtained when the restoring force produced by the system equals the force produced by the application of high external pressure.

Below I have explained this method for calculating energy of the system and the procedure to estimate the distortion for a given applied pressure.

2.1.1 Computing molecular energies using quantum theory

Solving a quantum chemical problem starts with finding a solution for time independent Schrödinger equation using the molecular Hamiltonian. This solutions gives an insight into the electronic structure of the molecule or in other words its chemical properties. Only hydrogen atom and hydrogen like systems have an exact solution for Schrödinger equation [4] because of the presence of only two particles (one nucleus and one electron), which makes it an effective one-particle problem [5]. Since all other atoms or molecules have more than two particles, their corresponding Schrödinger equations cannot be solved exactly and so approximate solutions are required for performing quantum calculations for a system other than hydrogen atom like systems.

The following equations give the Schrödinger equation and its corresponding Hamiltonian for a system of electrons and nuclei.

$$\hat{H}\psi = E\psi$$

$$\hat{H} = \hat{T}_{el} + V_{el-el} + \hat{T}_{nu} + V_{nu-nu} + V_{el-nu}$$

The exact Hamiltonian is given by $\hat{H}$, whereas $\hat{T}_{el}$ and $\hat{T}_{nu}$ are the kinetic energies of electrons and nuclei respectively, V_{el-el} and V_{nu-nu} are the potential energies of electrons and nuclei respectively, V_{el-nu} is the interaction between electrons and nuclei.

The Schrödinger equation is solved approximately using various techniques which is the subject area of quantum chemistry. One of these approximations is including orbital approximation which is based on the idea of using one-electron wave-functions to describe many-electron systems.

In practice, solving the Schrödinger equation usually involves making the so called Born-Oppenheimer (BO) approximation into account [6]. According to the Born-Oppenheimer approximation the total wave-function of the system can be separated

into electronic and nuclear parts. The first step in BO approximation is to study the electronic behaviour of a system by “freezing” the nuclei. Because electrons are very fast compared to nuclei it is safe to assume that interactions related to them happen at different time scales, which does not warrant there independent motion. This approximation is very useful when dealing with ground electronic state.

At “frozen” nuclei configuration $\hat{T}_{nu}$ goes to zero and V_{nu-nu} becomes a constant. The remaining terms constitute the complete electronic Hamiltonian.

$$\hat{H}_{el} = \hat{T}_{el} + V_{el-el} + V_{el-nu}$$

Solving the Schrödinger equation for electrons using this electronic Hamiltonian ($\hat{H}_{el}$) gives an idea of electronic structure at that particular “frozen” nuclei configuration. Using BO approximation gives potential energy surface as a function of inter-nuclear distances.

For this project the nuclei are frozen at various different geometries and the electronic Hamiltonian at this geometry is used to solve the Schrödinger equation. The approximate method used to solve the Schrödinger equation is second order Møller–Plesset perturbation theory (MP2) [7]. MP2 theory is one of several quantum chemistry post-Hartree–Fock ab initio methods, which improves on the Hartree – Fock method by adding electron correlation effects. The accuracy of this method also depends on the basis set chosen. The bigger the basis set the higher the accuracy of the results. The basis set used for these calculations is 6-31G (d, p) ++ basis set [8].

2.1.2 Procedure to estimate distortions due to external pressure

The DNA model formulated for this project is a 12 base pair B-DNA taken from the crystal structure from Research Collaboratory for Structural Bioinformatics (RCSB) protein data bank with structure ID ‘1BNA’ and sequence (5'-

D(*CP*GP*CP*GP*AP*AP*TP*TP*CP*GP*CP*G)-3') [9]. One pitch of a general B-DNA contains approximately ten nucleotides within a height of 3.4 nm (or 34 Å) [10].

For quantum calculations the effect of water solvent is ignored because of computational reasons. On applying external hydrostatic pressure on the DNA, it is assumed that all the covalent bonds (hard bonds) will not be affected and only the hydrogen bonds (soft bonds) will be affected.

Because of the assumption that the external hydrostatic pressure applied on the DNA system will affect the base pairs of the nucleotides in such a way that only the hydrogen bonds are affected, there is a need to determine the effective area on which the external hydrostatic pressure should be applied to get the desired changes in the hydrogen bonds. This effective area of the rectangle where the pressure is applied can be pictured as shown in Figure 2. This rectangle differs for A-T and G-C base pairs. For A-T base pair this rectangle has a height of 3.4 Å and its width of 5.7 Å completely includes the length of the base pair as shown in Figure 2 (b) and similarly for G-C base pair the rectangle has a height of 3.4 Å and a width of 6.5 Å.

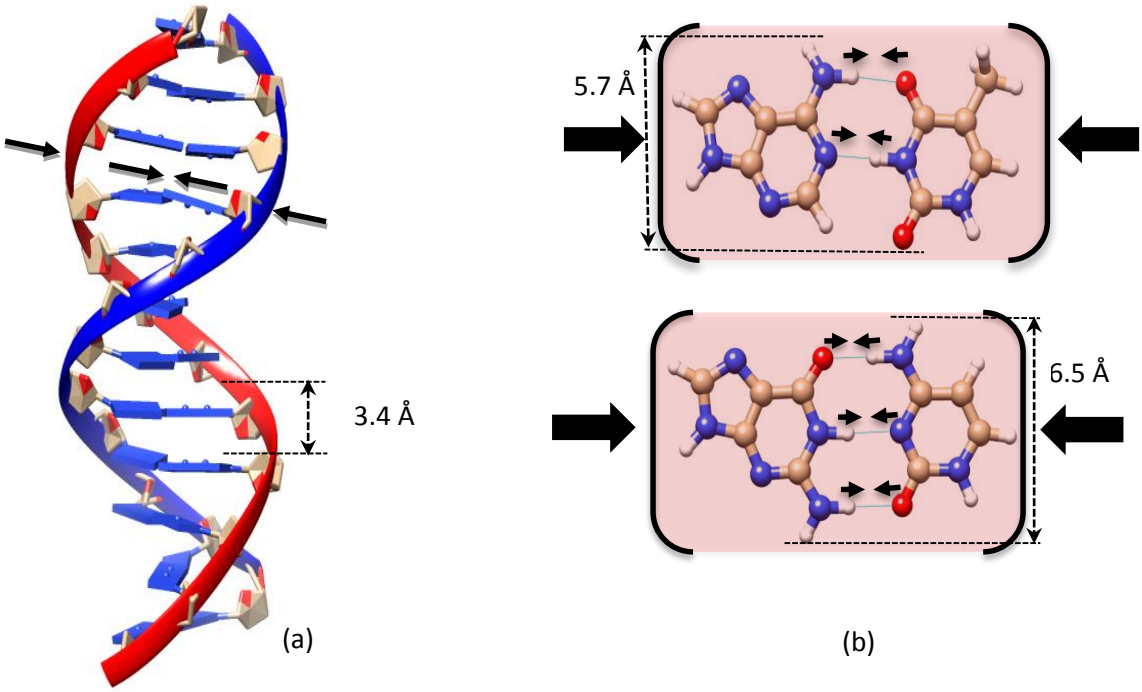


Figure 2: (a) B-DNA model. Force due to hydrostatic pressure is applied along the hydrogen bonds on the base pairs indicated by arrows. (b) The side brackets represent the area where external hydrostatic pressure is effective for hydrogen bond compressions on the base pairs of the DNA.

Using MP2 level of theory and the 6-31++G (d, p) basis set, the optimal energy structure of the individual bases and their corresponding base pairs were obtained. From these starting structures, displacements were made along the hydrogen bonds keeping the structure of the individual bases fixed and the corresponding single point energy was calculated. All calculations were performed using the GAMESS program package [11,12].

The energy of the displaced structure depends on six displacement internal coordinates ($\Delta x_1, \Delta x_2, \ldots \Delta x_6$) which determine the relative positions of the two bases, given that the structures of the individual bases are fixed. These six internal coordinates are the ones which are responsible for the connection between the two bases of the base pair. An easy way to understand this is explained in the following way: Each base of the base pairs will have its own set of internal coordinates. For e.g. in the case of A-T base pair suppose 'A' has ($3n$ – 6) internal coordinates and 'T' has ($3m$ – 6) internal coordinates but the base pair itself will have ($3(n + m)$ – 6) internal coordinates. There is an extra set of 6 internal coordinates which define the connection between 'A' and 'T'.

The energy can be fit to the following quadratic expression:

$$\Delta E = \frac{1}{2} \sum_{i,j=1}^{6} k_{ij} \Delta x_i \, \Delta x_j$$

This analytic expression for energy gives a many-to-one mapping of the six coordinates to energy, and a similar many-to-one mapping of the six coordinates to the restoring force on application of external pressure because there are several different structures possible for a given pressure. Clearly, this does not happen in the actual DNA because of various constraints due to its backbone and interactions with neighbouring base pairs and water molecules – there is a unique distorted structure for a given pressure.

To obtain a unique distorted structure for a given pressure, I have formulated the following procedure: which is based on the assumption that the base pairs when compressed translate and the displacements of the hydrogen bonds are done by translating one of the monomers along the axis which connects the pivot points between the two (almost parallel) hydrogen bonds as shown in Figure 3. This is done by using Cartesian coordinate system such that the required axis is made to be along one of the axes (say z axis) and then translate the monomer along this axis by changing only the (z) coordinate. This is done for the G-C base pair and the A-T base pair along their corresponding translational axes shown in Figure 3.

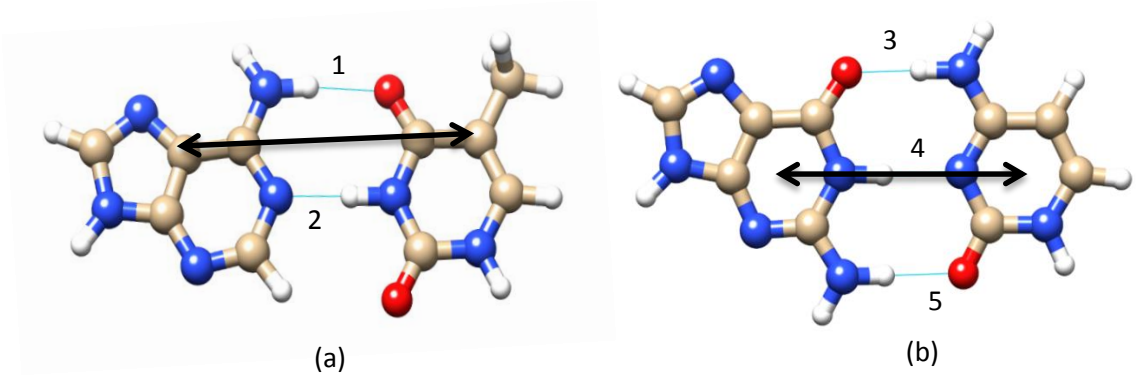


Figure 3: The double headed arrow is the translational axis connecting the pivot points used for translation in both (a) A-T and (b) G-C base pairs. (1, 2, 3, 4, and 5 are the labels for the respective hydrogen bonds).

This translation is then accompanied by the generation of a profile for the energy vs. displacement in the hydrogen bonds. The single point energy of these displaced dimers is calculated and plotted against the corresponding displacements. The data is then used to calculate the force using numerical differentiation keeping in mind that force is the negative derivative of energy. The force is then plotted against the corresponding displacements.

Using the force profiles for the individual base pairs and the known pressure, displacements were calculated for the applied pressure. This applied pressure is then converted into force by using the estimated area on which the pressure will be effective.

2.2 Molecular dynamics simulations

The molecular dynamics simulations are used in this study to obtain average equilibrium bond distances and angles. By performing molecular dynamics simulations at both normal pressure and high external hydrostatic pressure one can study the average changes in the hydrogen bond lengths and other observable structural changes in the B-DNA (like changes in grooves).

In the following discussion I have explained the theory and important points taken into consideration while performing these simulations.

2.2.1 Principles of molecular dynamics

Molecular dynamics (MD) is used to simulate the physical movements of atoms and molecules which govern microscopic and macroscopic properties and behaviours of the physical system. The interaction between these atoms and molecules is given by a defined potential. The trajectories of atoms and molecules are computed numerically by solving the Newton's equations of motion for the system of interacting particles. The forces and potential energies between these interacting particles are defined by the so called force fields.

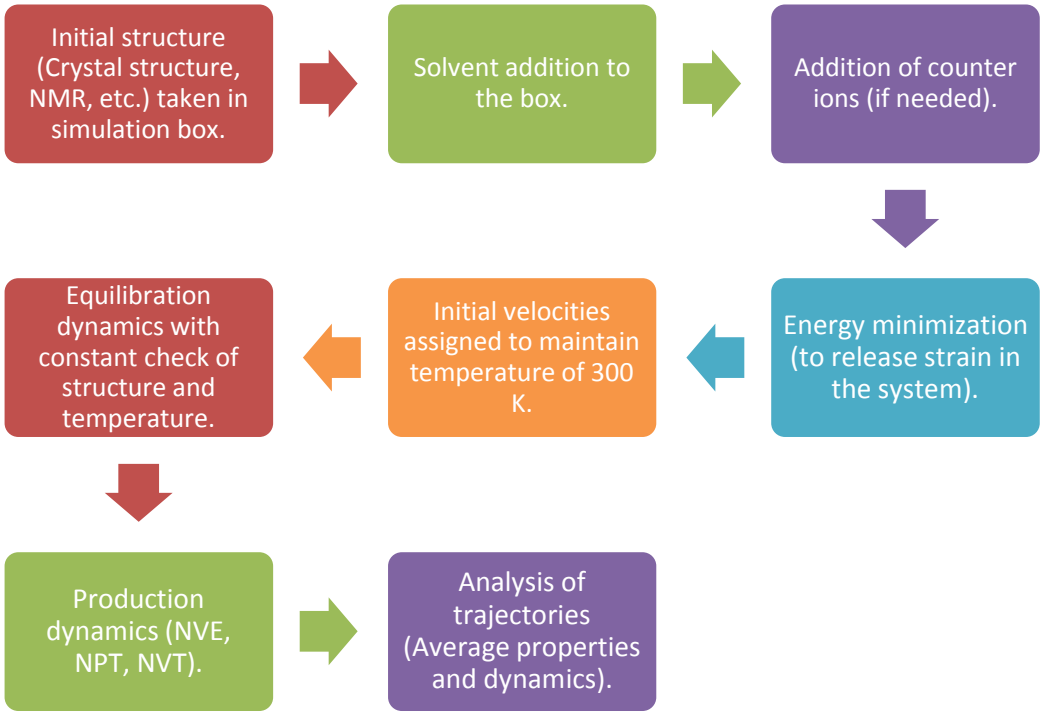
The force field gives information regarding the form and parameters of the functions which describe the potential energy of the system of particles [13]. The information which describes the potential energy of the system is the depth and width of the potential energy surface when plotted w.r.t inter particle distance. Force field functions and parameter sets are derived from both experimental work and quantum mechanical calculations. There are different types of force fields, for e.g. "All-atom" force fields which provides parameters for every type of atom in the system, "united-atom" force fields which does not include the representation of non-polar hydrogens explicitly and only the polar hydrogens are included, "coarse-grained" force fields are

used in long-time simulations of proteins and provide more crude representations to increase computational efficiency.

The water models are used to simulate water clusters, aqueous solutions, and liquid water with solvents. These water models contain information about the interaction centres, structural properties (like bond lengths, bonds angles, rigidity of the water molecule), electrostatic properties i.e. charge distribution, and also information about polarization effects [14].

Numerical methods are used in these simulations rather than the analytical methods because of the insolvable three-body problem analytically. The running time for the MD simulations also makes a difference as small runs do not give much information about the dynamics of the system and longer runs will accumulate errors in numerical integration. This part of accumulating errors can be minimized depending on the selection of algorithms and parameters but cannot be eliminated entirely.

A general MD simulation consists of different parts and is explained in the following flow chart:



Notes:

1. Counter ions are needed to imitate biological system by counteracting the charges present on the B-DNA (which for the model B-DNA is -22) and also to maintain physiological environment. For this study 150 mM of NaCl salt concentration was used to maintain the physiological condition [15].
2. Energy minimization is required to maintain a stress free environment in the system before providing initial velocities and heating up the system.
3. Initial velocities are randomly given based on the Maxwell-Boltzmann distribution to heat the system to a temperature of 300 K.
4. The equilibration process for a system with DNA molecule as a solute, consists of applying restraints on the solute and let the solvent settle (equilibrate around the solute). This part is continued with a little bit of decrease in the restraints applied on the solute. This allows a proper settling of solvent around solute and in case of DNA this allows proper settling of water in the major and minor grooves.
5. This equilibrated system is then used for simulations under the required parameters like for this study the equilibrated system is then simulated when 2000 bar of pressure is applied for a period of 1 ns, 2 ns and 2.5 ns giving different initial velocities for each simulation.

2.2.2 Specific simulation details

The technical details for the molecular dynamics simulations carried out in this study are given below:

The B-DNA model system is taken from the crystal structure from the RCSB protein data bank. Simulation are carried out with periodic boundary conditions within a cubic box using GROMACS program package with AMBER99sb force field [16].

The salt (NaCl) concentration used in the simulations is 150 mM and the number of ions is adjusted to ensure a zero net charge in the system. Counter ions are initially placed at random within the simulation box. The system is then solvated

with a layer of water 1.2 nm thick. Water is modelled using the TIP3P water model parameters [17].

The simulation involved 6,798 water molecules and 21,212 atoms in total. Electrostatic interactions are treated using the particle mesh Ewald method [18] with a real-space cut-off of 1 nm and cubic spline interpolation onto the charge grid with a spacing of 0.16 nm. The cut-off used for the Lennard-Jones interactions is 1 nm. Initial equilibration, involving energy minimization of the solvent, then of the solute-solvent system, followed by a slow heating. This protocol is described in ABC (Ascona B-DNA Consortium) publications [19].

Production simulations are carried out using an NPT ensemble and the Nose-Hoover algorithm [20] to control temperature and Parrinello-Rahman algorithm [21] to control pressure, with coupling constants of 0.4 ps for temperature and 1.0 ps for pressure. All chemical bonds involving hydrogen atoms are restrained using LINCS algorithm [22], allowing for stable simulations with a 0.5 fs time step.

The model is then simulated for 1 ns, 2 ns and 2.5 ns, saving trajectories at every 2 ps.

Chapter 3

Results and discussion

The following sections contains the results obtained from both quantum chemical calculations (section 3.1) and molecular dynamics simulations (section 3.2) and a brief discussion on the results.

3.1 Quantum chemical calculations

The energies differences are calculated for different geometries of the base pairs (A-T and G-C), as explained in the section 2.1.2, are plotted as energy profiles against the corresponding displacements. These energy differences corresponds to the energy needed to distort the base pairs from their original structure by the corresponding displacement. The x-axis corresponds to the displacement (in Å) and the y-axis corresponds to the energy difference (in Kcal/mol).

The force is calculated as the negative derivate of the energy w.r.t displacements, and plotted as the force profile. The x-axis corresponds to the displacement (in Å) and the y-axis corresponds to the force (in N).

These plots are then used as calibration curves to get the respective displacement for a given high external pressure.

The system's anharmonicity, as one goes farther from the optimized structure (minimum point in the energy profile represents the optimized structure), is visible from the following plots.

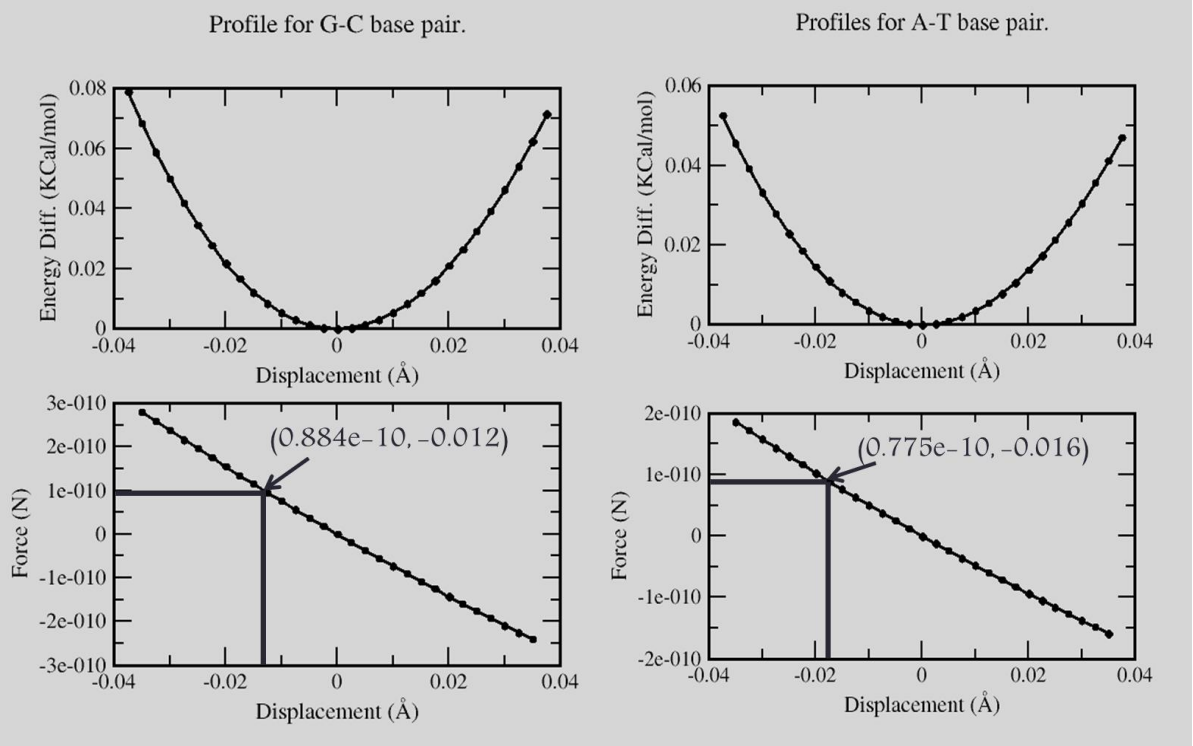


Figure 4: Energy and force profiles (vs. displacement) for both G-C and A-T base pair respectively.

- The force corresponding to 2000 bar pressure on A-T base pair is approx. 7.752×10^{-11} N and the corresponding displacement is approx. 0.016 Å.

- For G-C base pair:

  - The Force corresponding to 2000 bar pressure on G-C base pair is approx. 8.84×10^{-11} N and the corresponding displacement is approx. 0.012 Å.

The A-T base pair is compressed by 31.137% more than G-C base pair. This extra compression of the A-T base pair can be attributed to the presence of 3 hydrogen bonds in G-C base pair as compared to 2 hydrogen bonds in A-T, which is logical as

extra force is needed on G-C base pair to compress it by the same amount as the A-T base pair.

These quantum chemical calculations, although carried out with a fairly accurate method and a large enough basis set, suffers from some limitations inherent to the model used in this study.

The quantum calculations which have been carried out during this study are based on assumptions which do not fully represent the natural system. Not including the backbone and the water molecules in the model system are such examples. These exclusions were necessary and made to reduce the size of the system, which is a major limitation while performing quantum calculations. The difficulty in isolating the individual hydrogen bonds of the base pairs without modifying the entire base pair is also another limitation. Because of this limitation it was necessary to consider the average movement of the hydrogen bonds and for which pivot system was used to account for this average movement. Another limitation in performing quantum calculations is not being able to sample more in the configuration space and also not being able to include entropic effects which can be easily done in molecular dynamics simulations.

3.2 Molecular dynamics simulations

The B-DNA used for this study contains 8 G-C base pairs and 4 A-T base pairs. So, there are a total of 32 individual hydrogen bonds in this 12 base pair B-DNA (24 hydrogen bonds from G-C base pairs and 8 hydrogen bonds from A-T base pairs).

There are 5 types of hydrogen bonds among the base pairs, G-C base pair has 3 hydrogen bonds and A-T base pair has 2 hydrogen bonds. The labels used in for these 5 types of hydrogen bonds are AT (NH--O), AT (N--HN), GC (O--HN), GC (NH--N), and GC (NH--O) which correspond to 1, 2, 3, 4, and 5 hydrogen bonds of Figure 3 respectively. The 12 base pairs of the model B-DNA contains a total of 32 hydrogen bonds. GC (O--HN), GC (NH--N), and GC (NH--O) types have 8 hydrogen bonds each, whereas AT (NH--O), and AT (N--HN) types have 4 hydrogen bonds each. The

following plots show displacements of these 32 hydrogen bonds on applying high external hydrostatic pressure.

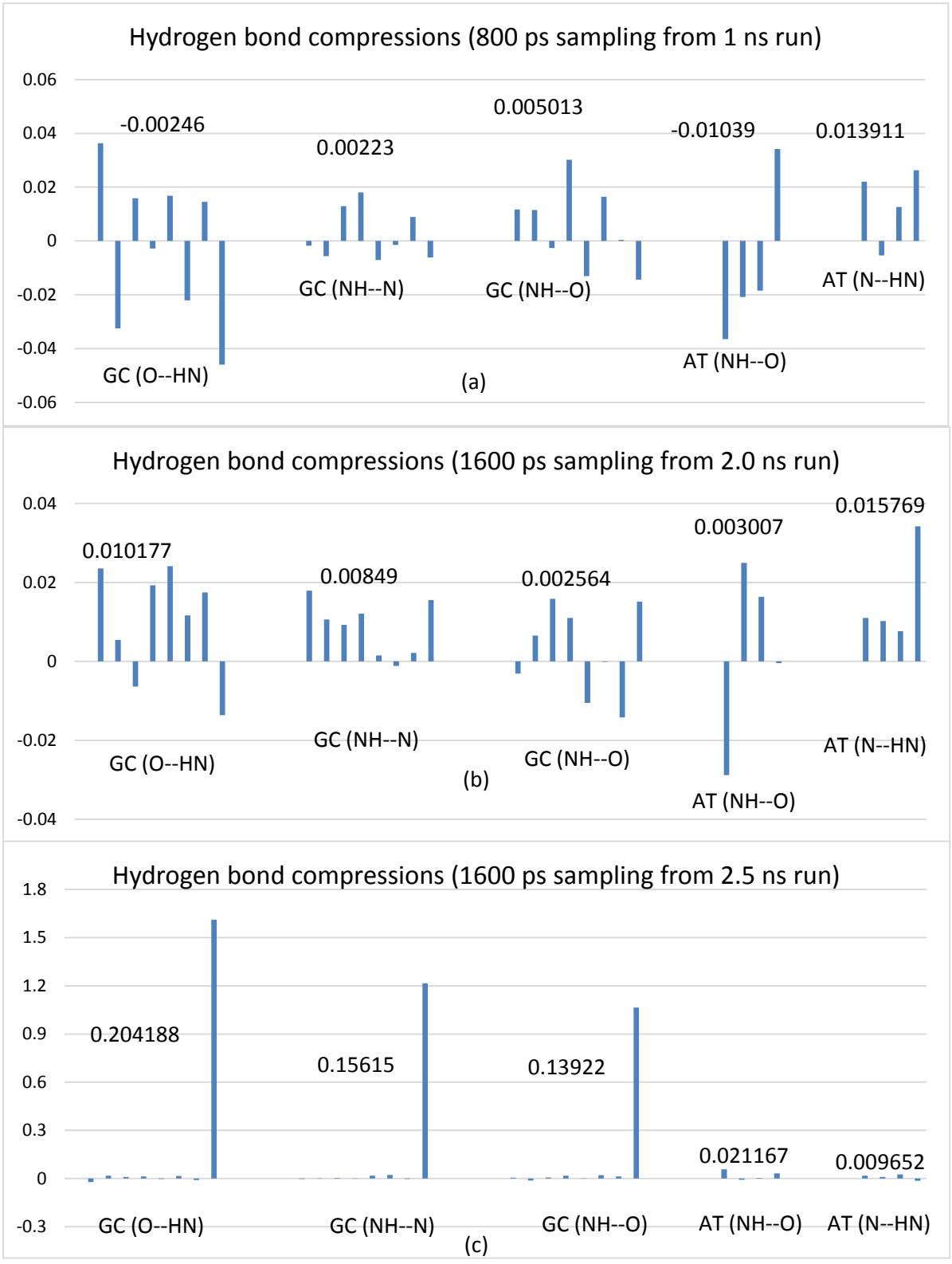


Figure 5: The above plots gives a histogram representation and the average values of hydrogen bond compressions. (a) Sampling time of 800 ps taken from 1 ns simulation.

(b) Sampling period of 1600 ps taken from 2 ns simulation and (c) Sampling period of 1600 ps taken from 2.5 ns simulation.

Going from left to right in the following plots, the displacements in the G-C hydrogen bonds are shown first followed by the A-T hydrogen bonds. The three types of G-C hydrogen bonds and two types of A-T hydrogen bonds are clustered together, thus giving five clusters of data. Within each cluster, the order of the bonds is according to sequence in which they appear in the B-DNA sequence starting from the 5’ end to the 3’ end of the B-DNA.

The results obtained from these molecular dynamics simulations indicate a trend corresponding to compressions of the hydrogen bonds in the base pairs of the B-DNA. However, given that there is not much correlation between the three different simulations, further investigation is required before arriving at definite conclusions.

Chapter 4

Conclusion

Quantum chemical calculations and molecular dynamics simulations suggests that there the DNA witnesses compressions in its base pairs when subjected to high external hydrostatic pressure. The compressions, calculated using quantum methods, for A-T base pair is 0.016 Å and for G-C base pair it is 0.012 Å.

But there is a discrepancy between experimental and theoretical results comes on comparing the actual displacements. The reported compressions (experiments) for the A-T base pair is 0.29 Å and for G-C base pair it is 0.11 Å [1].

One possible explanation for this is the assumptions taken into consideration while creating the model system. Another possible explanation for this discrepancy can be attributed to the misinterpretation of the relation between ^{1}H NMR chemical shifts and the hydrogen bond lengths of the DNA in the reported experimental results.

The relation between NMR chemical shifts with hydrogen bond lengths were derived for proteins [2, 3]. These may not directly transfer over to the DNA because of the fundamental difference between the hydrogen bonds of DNA and those of the proteins. The hydrogen bonds in proteins are formed among the backbone of the protein, whereas in DNA the hydrogen bonds are formed among the bases (i.e. base pairing). Further molecular dynamics simulations are required to obtain definitive conclusions.

Bibliography

[1] Wilton, David J., Mahua Ghosh, K. V. A. Chary, Kazuyuki Akasaka, and Mike P. Williamson. 2008. “Structural change in a B-DNA helix with.” *Nucleic Acids Research* 36 (12): 4032 - 4037. doi:10.1093/nar/gkn350.

[2] Williamson, M.P, and T. Asakura. 1993. “Empirical comparisons of models for chemical shift calculation in proteins.” *J. Magn. Reson. Ser. B* 101 (1): 63-71. doi:10.1006/jmrb.1993.1008.

[3] Asakura, T., K. Taoka, M. Demura, and M.P. Williamson. 1995. “The relationship between amide proton chemical shifts and secondary structure in proteins.” *J. Biomol. NMR* 6 (3): 227-236. doi:10.1007/BF00197804

[4] Schrödinger, Erwin. 1926. “An Undulatory Theory of the Mechanics of Atoms and Molecules.” *The Physical Review* 1049-1070.

[5] Szabo, Attila, and Ostlund, Neil S. “Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.” Courier Dover Publications, 1996

[6] Born, Max., J.R. Oppenheimer, and Translation by S.M Blinder. 1927. “Zur Quantentheorie der Molekeln [On the Quantum Theory of Molecules].” *Annalen der Physik (in German)* 389 (20): 457-484. doi:10.1002/andp.19273892002.

[7] Head-Gordon, Martin, and John A. Pople. 1988. “MP2 energy evaluation by direct methods.” *Chemical Physics Letters* 153 (6): 503-506. doi:10.1016/0009-2614(88)85250-3.

[8] Hariharan, P.C., and J.A. Pople. 1973. “The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies.” *Theoret. chim. Acta* 28 (3): 213-222. doi:10.1007/BF00533485.

[9] Drew, H.R., R.M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R.E. Dickerson. 1981. “Structure of a B-DNA dodecamer: Conformation and Dynamics.” *Proc. Natl. Acad. Sci. USA* 78 (4): 2179-2183. doi:10.2210/pdb1bna/pdb.

[10] Watson, J.D., and F.H.C. Crick. 1953. “Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acids.” *Nature* 171 (4356): 737. doi:10.1038/171737a0.

[11] Schmidt, M.W., K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, et al. 1993. “General Atomic and Molecular Electronic Structure System.” *J. Comput. Chem.* 14 (11): 1347-1363. doi:10.1002/jcc.540141112.

[12] Gordon, M.S., and M.W. Schmidt. 2005. “Advances in electronic structure theory: GAMESS a decade later.” In *Theory and Applications of Computational Chemistry: the first forty years*, edited by C.E. Dykstra, G. Frenking, K.S. Kim and G.E. Scuseria, 1167-1189. Amsterdam: Elsevier.

[13] *Force Field (Chemistry).* http://en.wikipedia.org/wiki/Force_field_(chemistry).

[14] *Water Model.* http://en.wikipedia.org/wiki/Water_model.

[15] Dang, Liem X. 1995. “Mechanism and Thermodynamics of Ion Selectivity in Aqueous Solutions of 18-Crown-6 Ether: A Molecular Dynamics Study.” *J. Am. Chem. Soc.* 117 (26): 6954–6960. doi:10.1021/ja00131a018.

[16] Hornak, Viktor, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. 2006. “Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters.” *Proteins* 712-725. doi:10.1002/prot.21123.

[17] Mahoney, Michael W., and William L. Jorgensen. 2000. “A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions.” *J. Chem. Phys.* 112 (20): 8910-8922. doi:10.1063/1.481505.

[18] Essmann, Ulrich, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. 1995. “A smooth particle mesh Ewald method.” *J. Chem. Phys.* 103 (19): 8577-8593. doi:10.1063/1.470117.

[19] *Ascona B-DNA Consortium: Publications.* http://gbiopbil.ibcp.fr/ABC/Publications.html.

[20] Hoover, William G. 1985. “Canonical dynamics: Equilibrium phase-space distributions.” *Phys. Rev. A* 31 (3): 1695-1697.

[21] Parrinello, M., and A. Rahman. 1981. “Polymorphic transitions in single crystals: A new molecular dynamics method.” *J. Appl. Phys.* 52 (12): 7182-7190. doi:10.1063/1.328693.

[22] Hess, Berk, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. 1997. “LINCS: A linear constraint solver for molecular simulations.” *J. Comp. Chem.* 18 (12): 1463-1472. doi:10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

.